

#7

1652

y. Pak

KUSH

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/462,845

DATE: 09/24/2001

TIME: 12:08:48

Input Set : A:\GC382-US-seqlist.txt

Output Set: N:\CRF3\09242001\I462845.raw

```

3 <110> APPLICANT: Estell, David A.
5 <120> TITLE OF INVENTION: Proteases From Gram-Positive Organisms
7 <130> FILE REFERENCE: GC382-US
9 <140> CURRENT APPLICATION NUMBER: US 09/462,845
10 <141> CURRENT FILING DATE: 2000-01-13
12 <150> PRIOR APPLICATION NUMBER: PCT/US98/14647
13 <151> PRIOR FILING DATE: 1998-07-14
15 <150> PRIOR APPLICATION NUMBER: EP 97305237.7
16 <151> PRIOR FILING DATE: 1998-07-15
18 <160> NUMBER OF SEQ ID NOS: 10
20 <170> SOFTWARE: FastSEQ for Windows Version 3.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1971
24 <212> TYPE: DNA
25 <213> ORGANISM: Bacillus subtilis
27 <400> SEQUENCE: 1
28 atgaaaaagc tgataaccgc agacgacatc acagcgattg tctctgtgac cgatcctcaa      60
29 tacgccccag acggtaccgc tgccgcatat gtaaaatcac aagtaaataca agagaaagat      120
30 tcgtatacat caaatatatg gatctatgaa acgaaaacgg gaggatctgt tccttgga      180
31 catggagaaa agcgaagcac gcacccaaga tggctctcgg acgggcgcac gcttgccctt      240
32 atttctgacg gagaaggcga tgcggcacag ctttatatca tgagcactga agcgaggaaa      300
33 gcaagaaaac tgactgatat cccatatggc gtgtcaaagc cgctatggc cccggacggc      360
34 gaatcgattc tggtcactat cagtttgagg gagggggaaa gcattgatga ccgagaaaaa      420
35 acagagcagg acagctatga acctgttgaa gtgcaaggcc tctcctacaa acgggacggc      480
36 aaagggctga cgagaggtgc gtatgccagc cttgtgcttg tcagcgtaaa gtcgggtgag      540
37 atgaaagagc tgacaagtca caaagctgat catggtgatc ctgctttttc tctgacggc      600
38 aaatggcttg ttttctcagc taatttaact gaaacagatg atgccagcaa gccgcatgat      660
39 gtttacataa tgtcactgga gtctggagat ctttaagcagg ttacacctca tcgcggtca      720
40 ttcggatcaa gctcattttc accagacgga aggtatcttg ctttgcttgg aaatgaaaag      780
41 gaatataaga atgctacgct ctcaaaggcg tggctctatg atatcgaaca aggccgcctc      840
42 acatgtctta ctgagatgct ggacgttcat ttagcggatg cgctgattgg agattcattg      900
43 atcgggtggtg ctgaacagcg cccgatttgg acaaaggaca gccaaagggtt ttatgtcatc      960
44 ggcacagatc aaggcagtag gggcatctat tatatttcga ttgaaggcct tgtgtatccg      1020
45 attcgtcttg aaaaagagta catcaatagc ttttctcttt cacctgatga acagcacttt      1080
46 attgccagtg tgacaaagcc ggacagaccg agtgagcttt acagtatccc gcttggaacag      1140
47 gaagagaaac agctgactgg cgcgaatgac aagtttgtca gggagcatac gatatcaata      1200
48 cctgaagaga ttcaatatgc tacagaagac ggcgtgatgg tgaacggctg gctgatgagg      1260
49 cctgcacaaa tgggaaggtga gacaacatat ccactttatc ttaacataca cggcggtccg      1320
50 catatgatgt acggacatac atattttcat gagtttcagg tgctggcggc gaaaggatac      1380
51 gcggtcgttt atatcaatcc gagaggaagc cacggctacg ggcaggaatt tgtgaatgcg      1440
52 gtcagaggag attatggggg aaaggattat gacgatgtga tgcaggctgt ggatgaggct      1500
53 atcaaacgag atccgcatac tgatcctaag cggctcgggtg tcacgggcgg aagctacgga      1560
54 ggtttttatga ccaactggat cgtcgggcag acgaaccgct ttaaagctgc cgttaccag      1620
55 cgctcgatat caaattggat cagctttcac ggcgtcagtg atatcggtta tttctttaca      1680
56 gactggcagc ttgagcatga catgtttgag gacacagaaa agctctggga ccggtctcct      1740
57 ttaaaatacg cagcaaacgt ggagacaccg cttttgatac tgcattggcg gcgggatgac      1800
58 cgatgcccga tcgagcaggc ggagcagctg tttatcgctc tgaaaaaaat gggcaaggaa      1860

```

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/462,845

DATE: 09/24/2001

TIME: 12:08:48

Input Set : A:\GC382-US-seqlist.txt

Output Set: N:\CRF3\09242001\I462845.raw

```

59 accaagcttg tccgttttcc gaatgcatcg cacaatttat cagcacccgg acacccaaga 1920
60 cagcggatca agcgcctgaa ttatatcagc tcatggtttg atcaacatct c 1971
62 <210> SEQ ID NO: 2
63 <211> LENGTH: 657
64 <212> TYPE: PRT
65 <213> ORGANISM: Bacillus subtilis
67 <400> SEQUENCE: 2
68 Met Lys Lys Leu Ile Thr Ala Asp Asp Ile Thr Ala Ile Val Ser Val
69 1 5 10 15
70 Thr Asp Pro Gln Tyr Ala Pro Asp Gly Thr Arg Ala Ala Tyr Val Lys
71 20 25 30
72 Ser Gln Val Asn Gln Glu Lys Asp Ser Tyr Thr Ser Asn Ile Trp Ile
73 35 40 45
74 Tyr Glu Thr Lys Thr Gly Gly Ser Val Pro Trp Thr His Gly Glu Lys
75 50 55 60
76 Arg Ser Thr Asp Pro Arg Trp Ser Pro Asp Gly Arg Thr Leu Ala Phe
77 65 70 75 80
78 Ile Ser Asp Arg Glu Gly Asp Ala Ala Gln Leu Tyr Ile Met Ser Thr
79 85 90 95
80 Glu Gly Gly Glu Ala Arg Lys Leu Thr Asp Ile Pro Tyr Gly Val Ser
81 100 105 110
82 Lys Pro Leu Trp Ser Pro Asp Gly Glu Ser Ile Leu Val Thr Ile Ser
83 115 120 125
84 Leu Gly Glu Gly Glu Ser Ile Asp Asp Arg Glu Lys Thr Glu Gln Asp
85 130 135 140
86 Ser Tyr Glu Pro Val Glu Val Gln Gly Leu Ser Tyr Lys Arg Asp Gly
87 145 150 155 160
88 Lys Gly Leu Thr Arg Gly Ala Tyr Ala Gln Leu Val Leu Val Ser Val
89 165 170 175
90 Lys Ser Gly Glu Met Lys Glu Leu Thr Ser His Lys Ala Asp His Gly
91 180 185 190
92 Asp Pro Ala Phe Ser Pro Asp Gly Lys Trp Leu Val Phe Ser Ala Asn
93 195 200 205
94 Leu Thr Glu Thr Asp Asp Ala Ser Lys Pro His Asp Val Tyr Ile Met
95 210 215 220
96 Ser Leu Glu Ser Gly Asp Leu Lys Gln Val Thr Pro His Arg Gly Ser
97 225 230 235 240
98 Phe Gly Ser Ser Ser Phe Ser Pro Asp Gly Arg Tyr Leu Ala Leu Leu
99 245 250 255
100 Gly Asn Glu Lys Glu Tyr Lys Asn Ala Thr Leu Ser Lys Ala Trp Leu
101 260 265 270
102 Tyr Asp Ile Glu Gln Gly Arg Leu Thr Cys Leu Thr Glu Met Leu Asp
103 275 280 285
104 Val His Leu Ala Asp Ala Leu Ile Gly Asp Ser Leu Ile Gly Gly Ala
105 290 295 300
106 Glu Gln Arg Pro Ile Trp Thr Lys Asp Ser Gln Gly Phe Tyr Val Ile
107 305 310 315 320
108 Gly Thr Asp Gln Gly Ser Thr Gly Ile Tyr Tyr Ile Ser Ile Glu Gly
109 325 330 335

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/462,845

DATE: 09/24/2001

TIME: 12:08:48

Input Set : A:\GC382-US-seqlist.txt

Output Set: N:\CRF3\09242001\I462845.raw

```

110 Leu Val Tyr Pro Ile Arg Leu Glu Lys Glu Tyr Ile Asn Ser Phe Ser
111           340           345           350
112 Leu Ser Pro Asp Glu Gln His Phe Ile Ala Ser Val Thr Lys Pro Asp
113           355           360           365
114 Arg Pro Ser Glu Leu Tyr Ser Ile Pro Leu Gly Gln Glu Lys Gln
115           370           375           380
116 Leu Thr Gly Ala Asn Asp Lys Phe Val Arg Glu His Thr Ile Ser Ile
117           385           390           395           400
118 Pro Glu Glu Ile Gln Tyr Ala Thr Glu Asp Gly Val Met Val Asn Gly
119           405           410           415
120 Trp Leu Met Arg Pro Ala Gln Met Glu Gly Glu Thr Thr Tyr Pro Leu
121           420           425           430
122 Ile Leu Asn Ile His Gly Gly Pro His Met Met Tyr Gly His Thr Tyr
123           435           440           445
124 Phe His Glu Phe Gln Val Leu Ala Ala Lys Gly Tyr Ala Val Val Tyr
125           450           455           460
126 Ile Asn Pro Arg Gly Ser His Gly Tyr Gly Gln Glu Phe Val Asn Ala
127           465           470           475           480
128 Val Arg Gly Asp Tyr Gly Gly Lys Asp Tyr Asp Asp Val Met Gln Ala
129           485           490           495
130 Val Asp Glu Ala Ile Lys Arg Asp Pro His Ile Asp Pro Lys Arg Leu
131           500           505           510
132 Gly Val Thr Gly Gly Ser Tyr Gly Gly Phe Met Thr Asn Trp Ile Val
133           515           520           525
134 Gly Gln Thr Asn Arg Phe Lys Ala Ala Val Thr Gln Arg Ser Ile Ser
135           530           535           540
136 Asn Trp Ile Ser Phe His Gly Val Ser Asp Ile Gly Tyr Phe Phe Thr
137           545           550           555           560
138 Asp Trp Gln Leu Glu His Asp Met Phe Glu Asp Thr Glu Lys Leu Trp
139           565           570           575
140 Asp Arg Ser Pro Leu Lys Tyr Ala Ala Asn Val Glu Thr Pro Leu Leu
141           580           585           590
142 Ile Leu His Gly Glu Arg Asp Asp Arg Cys Pro Ile Glu Gln Ala Glu
143           595           600           605
144 Gln Leu Phe Ile Ala Leu Lys Lys Met Gly Lys Glu Thr Lys Leu Val
145           610           615           620
146 Arg Phe Pro Asn Ala Ser His Asn Leu Ser Arg Thr Gly His Pro Arg
147           625           630           635           640
148 Gln Arg Ile Lys Arg Leu Asn Tyr Ile Ser Ser Trp Phe Asp Gln His
149           645           650           655
150 Leu
152 <210> SEQ ID NO: 3
153 <211> LENGTH: 818
154 <212> TYPE: PRT
155 <213> ORGANISM: Bacillus subtilis
157 <400> SEQUENCE: 3
158 Met Glu Gly Gly Glu Glu Glu Val Glu Arg Ile Pro Asp Glu Leu Phe
159   1           5           10           15
160 Asp Thr Lys Lys Lys His Leu Leu Asp Lys Leu Ile Arg Val Gly Ile

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/462,845

DATE: 09/24/2001

TIME: 12:08:48

Input Set : A:\GC382-US-seqlist.txt

Output Set: N:\CRF3\09242001\I462845.raw

```

161          20          25          30
162 Ile Leu Val Leu Leu Ile Trp Gly Thr Val Leu Leu Leu Lys Ser Ile
163          35          40          45
164 Pro His His Ser Asn Thr Pro Asp Tyr Gln Glu Pro Asn Ser Asn Tyr
165          50          55          60
166 Thr Asn Asp Gly Lys Leu Lys Val Ser Phe Ser Val Val Arg Asn Asn
167          65          70          75          80
168 Thr Phe Gln Pro Lys Tyr His Glu Leu Gln Trp Ile Ser Asp Asn Lys
169          85          90          95
170 Ile Glu Ser Asn Asp Leu Gly Leu Tyr Val Thr Phe Met Asn Asp Ser
171          100          105          110
172 Tyr Val Val Lys Ser Val Tyr Asp Asp Ser Tyr Asn Ser Val Leu Leu
173          115          120          125
174 Glu Gly Lys Thr Phe Ile His Asn Gly Gln Asn Leu Thr Val Glu Ser
175          130          135          140
176 Ile Thr Ala Ser Pro Asp Leu Lys Arg Leu Leu Ile Arg Thr Asn Ser
177          145          150          155          160
178 Val Gln Asn Trp Arg His Ser Thr Phe Gly Ser Tyr Phe Val Tyr Asp
179          165          170          175
180 Lys Ser Ser Ser Ser Phe Glu Glu Ile Gly Asn Glu Val Ala Leu Ala
181          180          185          190
182 Ile Trp Ser Pro Asn Ser Asn Asp Ile Ala Tyr Val Gln Asp Asn Asn
183          195          200          205
184 Ile Tyr Ile Tyr Ser Ala Ile Ser Lys Lys Thr Ile Arg Ala Val Thr
185          210          215          220
186 Asn Asp Gly Ser Ser Phe Leu Phe Asn Gly Lys Pro Asp Trp Val Tyr
187          225          230          235          240
188 Glu Glu Glu Val Phe Glu Asp Asp Lys Ala Ala Trp Trp Ser Pro Thr
189          245          250          255
190 Gly Asp Tyr Leu Ala Phe Leu Lys Ile Asp Glu Ser Glu Val Gly Glu
191          260          265          270
192 Phe Ile Ile Pro Tyr Tyr Val Gln Asp Glu Lys Asp Ile Tyr Pro Glu
193          275          280          285
194 Met Arg Ser Ile Lys Tyr Pro Lys Ser Gly Thr Pro Asn Pro His Ala
195          290          295          300
196 Glu Leu Trp Val Tyr Ser Met Lys Asp Gly Thr Ser Phe His Pro Arg
197          305          310          315          320
198 Ile Ser Gly Asn Lys Lys Asp Gly Ser Leu Leu Ile Thr Glu Val Thr
199          325          330          335
200 Trp Val Gly Asn Gly Asn Val Leu Val Lys Thr Thr Asp Arg Ser Ser
201          340          345          350
202 Asp Ile Leu Thr Val Phe Leu Ile Asp Thr Ile Ala Lys Thr Ser Asn
203          355          360          365
204 Val Val Arg Asn Glu Ser Ser Asn Gly Gly Trp Trp Glu Ile Thr His
205          370          375          380
206 Asn Thr Leu Phe Ile Pro Ala Asn Glu Thr Phe Asp Arg Pro His Asn
207          385          390          395          400
208 Gly Tyr Val Asp Ile Leu Pro Ile Gly Gly Tyr Asn His Leu Ala Tyr
209          405          410          415

```

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/462,845

TIME: 12:08:48

Input Set : A:\GC382-US-seqlist.txt

Output Set: N:\CRF3\09242001\I462845.raw

```

210 Phe Glu Asn Ser Asn Ser Ser His Tyr Lys Thr Leu Thr Glu Gly Lys
211           420           425           430
212 Trp Glu Val Val Asn Gly Pro Leu Ala Phe Asp Ser Met Glu Asn Arg
213           435           440           445
214 Leu Tyr Phe Ile Ser Thr Arg Lys Ser Ser Thr Glu Arg His Val Tyr
215           450           455           460
216 Tyr Ile Asp Leu Arg Ser Pro Asn Glu Ile Ile Glu Val Thr Asp Thr
217 465           470           475           480
218 Ser Glu Asp Gly Val Tyr Asp Val Ser Phe Ser Ser Gly Arg Arg Phe
219           485           490           495
220 Gly Leu Leu Thr Tyr Lys Gly Pro Lys Val Pro Tyr Gln Lys Ile Val
221           500           505           510
222 Asp Phe His Ser Arg Lys Ala Glu Lys Cys Asp Lys Gly Asn Val Leu
223           515           520           525
224 Gly Lys Ser Leu Tyr His Leu Glu Lys Asn Glu Val Leu Thr Lys Ile
225           530           535           540
226 Leu Glu Asp Tyr Ala Val Pro Arg Lys Ser Phe Arg Glu Leu Asn Leu
227 545           550           555           560
228 Gly Lys Asp Glu Phe Gly Lys Asp Ile Leu Val Asn Ser Tyr Glu Ile
229           565           570           575
230 Leu Pro Asn Asp Phe Asp Glu Thr Leu Ser Asp His Tyr Pro Val Phe
231           580           585           590
232 Phe Phe Ala Tyr Gly Gly Pro Asn Ser Gln Gln Val Val Lys Thr Phe
233           595           600           605
234 Ser Val Gly Phe Asn Glu Val Val Ala Ser Gln Leu Asn Ala Ile Val
235           610           615           620
236 Val Val Val Asp Gly Arg Gly Thr Gly Phe Lys Gly Gln Asp Phe Arg
237 625           630           635           640
238 Ser Leu Val Arg Asp Arg Leu Gly Asp Tyr Glu Ala Arg Asp Gln Ile
239           645           650           655
240 Ser Ala Ala Ser Leu Tyr Gly Ser Leu Thr Phe Val Asp Pro Gln Lys
241           660           665           670
242 Ile Ser Leu Phe Gly Trp Ser Tyr Gly Gly Tyr Leu Thr Lys Thr
243           675           680           685
244 Leu Glu Lys Asp Gly Gly Arg His Phe Lys Tyr Gly Met Ser Val Ala
245           690           695           700
246 Pro Val Thr Asp Trp Arg Phe Tyr Asp Ser Val Tyr Thr Glu Arg Tyr
247 705           710           715           720
248 Met His Thr Pro Gln Glu Asn Phe Asp Gly Tyr Val Glu Ser Ser Val
249           725           730           735
250 His Asn Val Thr Ala Leu Ala Gln Ala Asn Arg Phe Leu Leu Met His
251           740           745           750
252 Gly Thr Gly Asp Asp Asn Val His Phe Gln Asn Ser Leu Lys Phe Leu
253           755           760           765
254 Asp Leu Leu Asp Leu Asn Gly Val Glu Asn Tyr Asp Val His Val Phe
255           770           775           780
256 Pro Asp Ser Asp His Ser Ile Arg Tyr His Asn Ala Asn Val Ile Val
257 785           790           795           800
258 Phe Asp Lys Leu Leu Asp Trp Ala Lys Arg Ala Phe Asp Gly Gln Phe

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/462,845

DATE: 09/24/2001

TIME: 12:08:49

Input Set : A:\GC382-US-seq1ist.txt

Output Set: N:\CRF3\09242001\I462845.raw